

DOI:10.19651/j.cnki.emt.2416020

基于图像特征增强的低照度目标检测算法^{*}

黄玉龙 张晓玲

(江苏理工学院电气信息工程学院 常州 213001)

摘要: 低照度环境会导致图像目标特征不明显,噪声干扰严重等情况,影响目标检测器的检测性能。针对以上问题,构建了一个多尺度图像特征增强模块 FEM,并与 YOLOv8s 目标检测网络联合,构建了端到端的低照度目标检测算法 FE-YOLO。首先,使用 FEM 从输入图像中构建三个不同尺度下的特征信息并进行高效融合,得到具有丰富特征表达的增强图像。然后,在 YOLOv8s 颈部网络中添加目标特征增强模块 TFE,通过抑制高层特征中的背景噪声信息,突出目标特征的表达。实验结果表明:在低照度图像目标检测数据集 ExDark 上的平均精度均值 (mAP) 达到了 75.63%,与原始的 YOLOv8s 算法相比,提高了 3.03%,本文算法在低照度目标检测任务中取得了更好的检测效果。

关键词: 目标检测;低照度图像;YOLOv8 网络;图像增强;特征融合

中图分类号: TP391.4;TN60 **文献标识码:** A **国家标准学科分类代码:** 510.40

Low-light object detection algorithm based on image feature enhancement

Huang Yulong Zhang Xiaoling

(College of Electrical and Information Engineering, Jiangsu University of Technology, Changzhou 213001, China)

Abstract: Low illumination environments can lead to situations such as inconspicuous image target features and severe noise interference, which affect the detection performance of the object detector. To address the above problems, a multi-scale image feature enhancement module FEM is constructed, and in conjunction with YOLOv8s object detection network, an end-to-end low-light image object detection method FE-YOLO is constructed. Firstly, FEM is employed to extract feature information from the input image at three different scales and efficiently fuse them to obtain an enhanced image with rich feature representation. Then, in the neck network of YOLOv8s, a target feature enhancement module TFE is incorporated. TFE works by suppressing background noise information in higher-level features, thereby accentuating the representation capacity of target features. The experimental results show that the mean average precision mean (mAP) on the low-light image object detection dataset ExDark reaches 75.63%, which is 3.03% higher than the original YOLOv8s algorithm, and this paper's algorithm achieves a better detection result in the low-light object detection task.

Keywords: object detection; low-light images; YOLOv8 network; image enhancement; feature fusion

0 引言

基于深度学习的目标检测方法,近年来在许多通用目标检测数据集上取得了良好的性能。因此,目标检测算法得到了广泛应用。例如,在智能车辆环境感知、智能安防、行人检测和智慧城市等领域。实际应用时,在光照不足的环境下采集到的图像数据往往存在细节丢失、亮度低、噪点高等问题,导致通用目标检测算法难以提取到高质量的特征数据,进而影响检测精度。低照度目标检测器在夜间安

防、智慧交通、无人驾驶等场景具有很大的应用场景,因此,研究低照度环境下的目标检测技术具有实际意义。

近年来,随着计算机技术的发展,基于卷积神经网络(convolutional neural network, CNN)架构的模型在目标检测领域占据主要地位并取得了较高的检测性能^[1]。目前基于深度学习的目标检测模型按照其检测方法可分为两阶段检测模型与单阶段检测模型^[2]。两阶段检测模型从输入图像中生成若干区域,再使用卷积神经网络对每个区域进行特征提取以实现检测分类,例如 R-CNN^[3]、Fast R-CNN^[4]、

收稿日期:2024-05-12

^{*} 基金项目:国家自然科学基金(61305123)、产学研项目(KYH17134)资助

MaskRCNN^[5]等。两阶段检测模型能够有效地提取图像特征,检测精度较高,但对每个划分的区域进行一次卷积操作,计算量庞大,难以广泛应用。单阶段检测模型对图像区域进行划分,并采用多尺度特征模块一次性完成检测分类操作,例如 SSD^[6]、YOLO^[7]系列等。虽然检测精度相较于两阶段检测模型有部分降低,但检测速度更快,所需要的计算量更低,方便部署在更多场景下使用。

在正常光照数据集中研究的目标检测模型取得了良好的效果,并得到了广泛的应用。但在低照度环境中的目标检测模型研究较少。近几年,文献[8-10]采用对抗学习的方法将现有正常光线下的目标检测模型应用到暗光环境下。但是他们的工作重点是提取暗光图像中的特征数据去匹配正常光线下目标检测模型所需要的信息,而忽略了原始暗光图像所蕴含的潜在信息。这些方法需要预先在大批量数据集中对模型进行训练得到预训练模型。虽然可以提高检测器在暗光场景下的检测精度,但是这种方法需要消耗大量算力,不适合多种暗光环境中的部署。文献[11-13]提出一种可微分图像增强模块,通过设计一个图像处理模块与参数预估模块,参数预估模块自适应地学习一组参数指导图像处理模块对图像进行提高亮度、对比度、锐化等操作。通过图像处理模块增强后的图像质量得到了一定的提升,但是在对图像进行增强的同时会增加大量的噪声干扰信息,影响目标区域的特征信息表达。以上方法在提高图像质量的同时,忽略了噪声对检测精度的影响,并且图像增强模块设计复杂,导致算法的计算量较大。文献[14]通过知识蒸馏方法增强 YOLOv8 检测器的检测性能,并融合可见光图像与红外图像扩充训练数据,增强模型的训练效果。文献[15]使用 MobileNet v2^[16]替换了 YOLOv8 的骨干网络,增强模型的特征提取能力,并在颈部使用动态特征融合注意力块,增强不同层次特征的融合能力。文献[17]通过生成低照度图像,扩充数据集,增强模型的训练效果,并且提出周边信息模块,降低图像噪声的影响。上述这些方法通过扩充训练数据集以及改进目标检测器结构,增强特征提取与融合能力,得到了检测精度的提升,但没有充分挖掘原始低照度图像中目标区域的有效特征信息。

针对低照度环境下输入图像质量低、噪声干扰目标区域特征表达等问题,影响目标检测器的检测精度。本文提出一种适用于低照度环境下的目标检测算法, FE-YOLOv8。首先,本文建立了一个图像特征信息增强模块(feature enhancement module, FEM)通过融合不同尺度下的层次信息提高输入图像的目标特征表达。该模块与 YOLOv8 网络联合进行端到端训练,使用目标检测的损失优化模块参数。然后,在 YOLOv8 的颈部网络中引入目标特征增强模块(target feature enhancement, TFE),通过抑制背景噪声信息的干扰,增强目标区域的特征表达能力。实验结果表明,将本文所提算法应用于低照度目标检测,平均检测精度有较大提升。

1 相关工作

1.1 低照度环境下的目标检测

与一般的目标检测任务相比,从低照度的图像中检测目标仍然是一项具有挑战性的任务。除了引言中所提到的基于可微分图像增强模块^[11-13]的方法, Yin 等^[18]提出一种基于拉普拉斯金字塔的图像增强网络并与目标检测器连接,构成了低照度目标检测模型 PE-YOLO。该方法使用拉普拉斯金字塔对输入图像进行分解与重建,并在分解与重建过程中对不同尺度图像进行增强。但是分解重建过程容易受到图像中噪声的干扰,并且有可能在目标边缘区域产生偏差,影响输入图像的效果。Cui 等^[19]提出了一种多任务自动编码转换(multitask auto encoding transformation, MAET)模型,该模型通过预测退化参数并且对图像的退化变换进行编码和解码操作,将正常光线图像合成为暗光图像,并一起输入到网络中进行训练,提高训练效果。Feng 等^[20]提出了一种通过逆映射网络将输入图像转化为伪 RAW 图像,并增强伪 RAW 图像在机器视觉下的白平衡、色彩显示和亮度的方法,提高目标检测精度。上述两种方法通过设计复杂的图像处理模块,达到提高检测精度的效果。但是这种方法使网络结构变得更加复杂,算法计算量较大,不方便模型的大规模部署与使用。

1.2 YOLOv8 目标检测模型

YOLOv8 系列是性能出色的单阶段目标检测模型,应用在多个数据集上进行目标检测均能够取得很好的检测效果。YOLOv8 整体结构由主干网络、颈部结构和检测头构成。通过调整模型的宽度和深度,可以组合出 YOLOv8n、YOLOv8s、YOLOv8m、YOLOv8l、YOLOv8x 五种型号,这五种型号算法的检测精度依次提高,但算法所需计算量也依次增加,对设备算力的需求也逐渐增加。因此,综合考虑检测精度与算法计算量的平衡,本文在 YOLOv8s 模型的基础上进行改进。

2 本文算法

本文算法整体框架如图 1 所示。首先,将 FEM 与 YOLOv8 目标检测网络连接, FEM 是一个即插即用模块,通过级联 FEM 与 YOLOv8 目标检测网络进行端到端训练,并使用目标检测结果的损失对整体网络参数进行更新。FEM 通过融合图像的层次特征,自适应地增强输入图像的语义表示。其次,在 YOLOv8s 颈部网络的上采样之后添加 TFE 模块,通过减少背景噪声信息对目标特征的干扰,增强目标区域的特征表达,达到提升检测精度的效果。

2.1 FEM 模块结构

FEM 通过融合不同尺度的特征信息增强输入图像的细节信息与整体层次结构,有利于提高目标检测网络的检测精度。具体地说, FEM 首先使用 3 个卷积层处理输入图

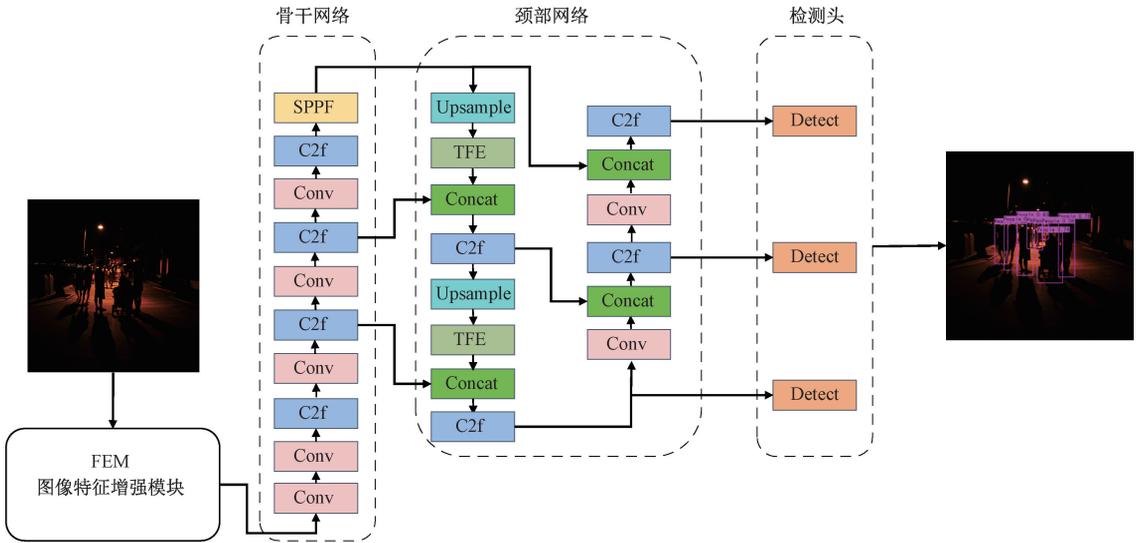


图 1 本文算法整体框架

Fig. 1 The overall framework of the algorithm in this paper

像,获得 3 个尺度的图像分量。然后将其输入特征增强网络(feature enhancement network, FENet),获得 3 个尺度

增强后的特征信息,最后将 3 个特征信息进行融合,获得最终的增强特征。FEM 结构如图 2 所示。

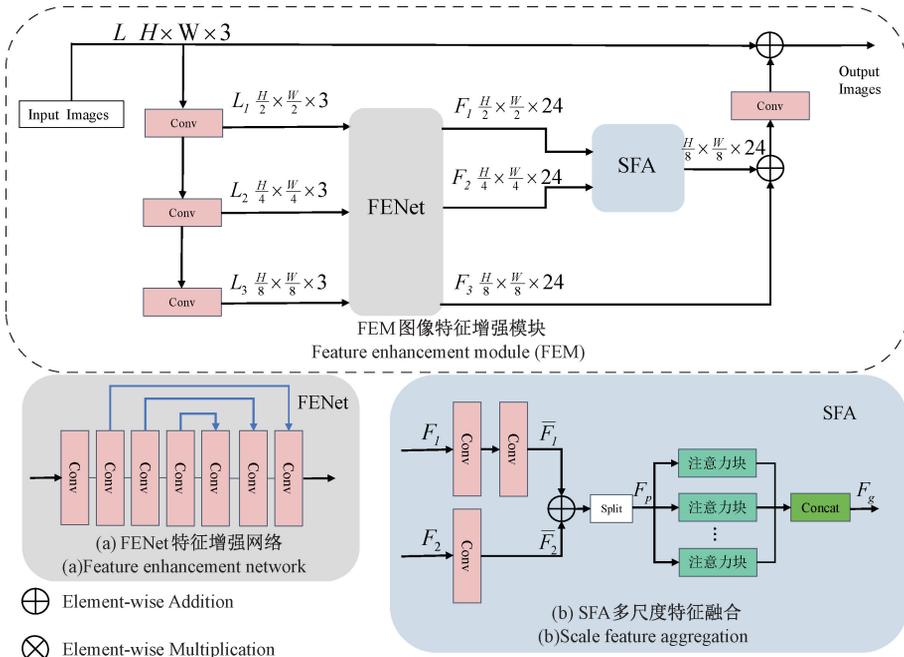


图 2 图像特征增强模块

Fig. 2 Image feature enhancement module

1) 构建多尺度图像

为了融合不同尺度下的层次特征,首先需要将输入的弱光图像处理为多尺度表示。然后,将这些多尺度的图像数据传输到特征增强网络,得到增强后的不同尺度下的特征信息。

本文将输入的弱光图像 $L \in R^{H \times W \times C}$ 作为输入,并使用普通卷积处理输入图像得到 3 个不同尺度下的图像 $L_1 \in$

$R^{\frac{H}{2} \times \frac{W}{2} \times 3}$ 和 $L_2 \in R^{\frac{H}{4} \times \frac{W}{4} \times 3}$ 以及 $L_3 \in R^{\frac{H}{8} \times \frac{W}{8} \times 3}$ 。分别表示原始输入图像宽度和高度的 1/2、1/4 以及 1/8。实现公式如下:

$$\begin{aligned}
 L_1 &= Conv(L) \quad K=7, S=2 \\
 L_2 &= Conv(L_1) \quad K=5, S=2 \\
 L_3 &= Conv(L_2) \quad K=3, S=2
 \end{aligned}
 \tag{1}$$

其中,Conv 表示普通卷积,K 和 S 表示卷积核的大小

和步幅, H 、 W 、 C 表示输入图像的高度, 宽度和通道数。

2) FENet 特征增强网络

受弱光图像增强的深度网络^[21]启发, 本文设计了一个全卷积的特征增强网络 FENet。网络结构如图 2(a) 所示。首先使用一个卷积生成特征图 $F \in R^{H \times W \times 32}$, 在保持输入图像宽高不变的前提下将通道 C 从 3 扩展为 32。然后使用 6 个对称跳跃连接的卷积层从输入图像中提取丰富的空间特征, 并增强层次特征的语义表示, 其中每个卷积层的卷积核大小 $K=3$, 步幅 $S=1$, 并在每个卷积层后使用 Relu 激活函数。使用对称跳跃连接的卷积层, 前面的卷积层提取到的特征可以作为后面卷积层的输入, 这样可以帮助网络更加全面地提取到不同卷积层所学习到的不同级别的特征信息, 从而增加了特征信息的多样性。与弱光图像增强的深度网络不同的是, 本文最后一个卷积层的输出为每个尺度最终增强后的特征表示, 而且, 为了保存相邻像素间的关系, 本文在 FENet 中舍弃了降采样与归一化处理。经 FENet 处理 L_1 、 L_2 和 L_3 得到增强后的多尺度特征 F_1 、 F_2 和 F_3 。通过这种方法可以增强 3 个尺度下图像中的全局信息和局部信息。同时可以保留相邻像素之间的语义关系。经过多尺度特征增强处理后, 最终获得增强后的 3 个尺度下特征表示。然后对多尺度特征 F_1 、 F_2 和 F_3 进行融合处理。

3) 多尺度特征融合

该部分主要实现融合 2.1 节 2) FENet 特征增强网络中获得的多尺度特征信息 F_1 、 F_2 和 F_3 。其中, 高尺度特征 F_1 是从较高分辨率图像中提取得到的, 高分辨率图像中的每个像素表示较小的区域, 因此, 从高分辨率图像中提取到的特征信息包含更多的抽象信息, 如图像的全局结构、目标形状等。而从低分辨率图像中提取到的尺度特征 F_3 则包含更多的细节信息以及边缘特征等。这是由于低分辨率图像中的每个像素代表着原始图像中较大的一片区域, 目标的一些局部信息和细微特征会被大的像素块表示, 这使得特征提取网络可以从低分辨率图像中更加高效地提取到目标的细节和边缘特征信息。由于简单地将不同尺度中的特征信息进行聚合不能高效地利用这些信息。因此, 受 Transformer 网络^[22]中多头注意力思想启发, 本文设计了一个尺度特征融合模块 (scale feature aggregation, SFA), 该模块通过从不同的通道中学习信息, 关注不同尺度下的特征, 达到高效融合特征的目的。具体结构如图 2(b) 所示, 在特征融合之前, 先将高尺度特征 F_1 、 F_2 处理到更低尺度。首先使用两个卷积层 ($K=5$, $S=2$; $K=3$, $S=2$) 将 $F_1 \in R^{\frac{H}{2} \times \frac{W}{2} \times 24}$ 转换为 $\bar{F}_1 \in R^{\frac{H}{8} \times \frac{W}{8} \times 24}$, 使用一个卷积层 ($K=3$, $S=2$) 将 $F_2 \in R^{\frac{H}{4} \times \frac{W}{4} \times 24}$ 转化为 $\bar{F}_2 \in R^{\frac{H}{8} \times \frac{W}{8} \times 24}$ 。并将 \bar{F}_1 和 \bar{F}_2 连接形成多层信息特征 F_p , 然后将 F_p 沿通道维度 C 分为 n 个块:

$$F_p^j = F_p[:, :, (j-1) \frac{C}{n} : j \frac{C}{n}] \quad (2)$$

式中: $j \in \{1, 2, \dots, n\}$, n 为注意力块的总数。

将多层信息特征沿着通道维度切分, 然后在每个子特征图上应用 Transformer 网络中的注意力块, 如图 3 所示。注意力块中包含两个顺序连接的子模块: 多头注意力块 (multi-dconv head transposed attention, MDTA) 和前馈网络模块 (gated-dconv feed forward network, GDFN)。多头注意力块的核心是自注意力机制, 自注意力机制从输入特征中得到查询向量 Q 、键向量 K 以及值向量 V 。其中查询向量 Q 表示输入特征中某一位置的特征表达, 键向量 K 表示输入特征中全部位置的特征表达。通过对查询向量 Q 和键向量 K 进行点积计算, 可以获得输入特征中某一位置相对于输入特征中其他位置的重要性。然后, 对点积计算结果进行归一化操作, 得到注意力权重。将切分后的多层信息特征 F_p^j 输入注意力块, 计算不同通道中的自注意力权重, 得到 F_m^j :

$$F_m^j = PC(V \cdot \sigma(K \cdot Q/\alpha)) + F_p^j \quad (3)$$

式中: 将输入特征归一化后使用 1×1 卷积和 3×3 卷积后得到 Q 、 K 、 V 。PC 表示点卷积 (pointwise convolution, PC), α 表示一组可学习的参数, (\cdot) 表示点积计算。

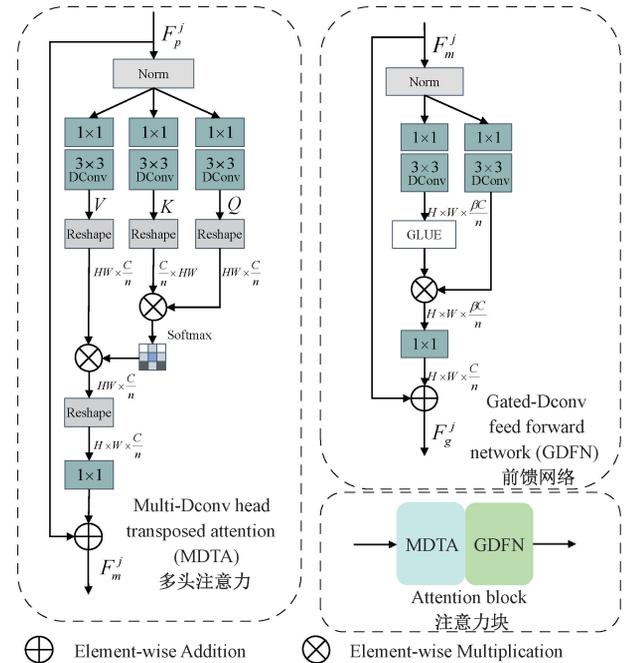


图 3 注意力块

Fig. 3 Attention block

然后, 将 F_m^j 输入前馈网络模块转化特征, 得到 F_g^j 。前馈网络通过对自注意力机制的计算结果引入非线性变换, 可以增强模型的表达能力, 前馈网络的计算公式如下:

$$\begin{aligned} X_1 &= \text{GELU}(\text{DC}_{3 \times 3}(\text{PC}(\text{LN}(F_m^j)))) \\ X_2 &= \text{DC}_{3 \times 3}(\text{PC}(\text{LN}(F_m^j))) \\ F_g^j &= \text{PC}(X_1 \times X_2) + F_m^j \end{aligned} \quad (4)$$

式中: GELU 表示激活函数, LN 是层归一化操作, DC 表

示深度可分离卷积。

最终,沿着通道维度将 n 个增强后的特征 F_g^j 拼接为最终输出特征 F_g

$$F_g = [F_g^1, \dots, F_g^j, \dots, F_g^n], j \in \{1, 2, \dots, n\} \quad (5)$$

通过这种方法可以加强不同通道间特征的交互作用,同时,对每个子特征图进行增强,使模型可以专注于学习该通道中的特征信息,实现更加高效的特征融合。因此,最终得到的特征 F_g 包含了更丰富的特征表示,包含了不同尺度下的信息。然后,使用双线性插值方法将低尺度下的特征 F_3 与融合后特征 F_g 进行上采样,将尺寸变换到与输入图像尺寸大小相同,然后融合特征 F_3 与特征 F_g 获得最终的增强特征。由于融合了低尺度特征中的细节信息以及高尺度特征中的抽象信息,最终可以获得更加全面和准确的特征信息,这增强了输入图像的细节表示以及全局信息特征,提高了目标检测网络的检测精度。

2.2 TFE 模块

受照明不足的影响,低照度图像中的目标往往与背景的背景区域混杂在一起。导致网络无法提取到显著的目标特征信息^[23-24]。随着网络深度的增加,使用更多的卷积处理特征信息,会导致目标特征信息与黑暗区域的背景噪声融合^[25],干扰检测器正确识别目标。为解决此问题,本文设计了 TFE 模块增强目标特征,如图 4 所示,通过利用目标的邻域信息,抑制背景噪声的干扰。具体来说,考虑到目标信息的各向异性与背景信息的自相似性,本文引入中心像素差值卷积(central pixel difference convolution, CPDC)^[26]来抑制背景噪声的影响。

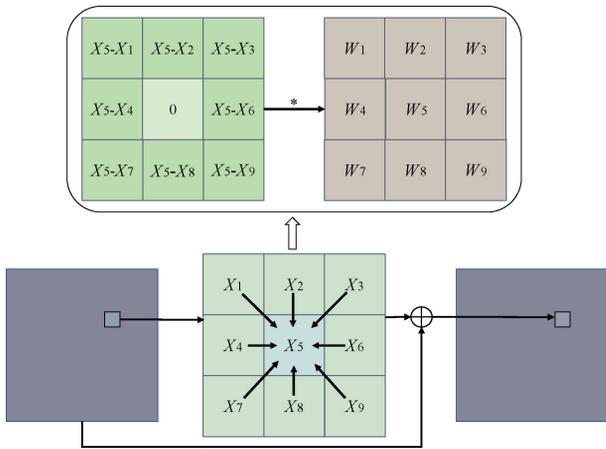


图 4 TFE 目标特征增强模块

Fig. 4 Target feature enhancement module

中心像素差值卷积的计算过程与普通卷积类似。其中,特征图中的原始像素被卷积核覆盖,并在卷积操作中被像素差值取代。普通卷积与中心像素差值卷积的计算公式可以表示为:

$$y = f(x, \theta) = \sum_{i=1}^{k \times k} w_i \cdot x_i, (\text{Vanilla conv}) \quad (6)$$

$$y = f(\nabla x, \theta) = \sum_{(x_i, x'_i) \in p} w_i \cdot (x_i - x'_i), (\text{CPDC}) \quad (7)$$

式中: $p = \{(x_1, x'_1), (x_2, x'_2), \dots, (x_m, x'_m)\}$ 是从特征图的局部块中选取的像素对集合, x_i 和 x'_i 表示输入像素, w_i 表示 $k \times k$ 卷积核中的权重。其中, $m \leq k \times k$ 。

中心像素差值卷积的计算原理是通过计算局部窗口中的中心像素与其周围像素之间的差值,然后基于所得差值生成新的特征图。通过计算中心像素与其周围像素之间的差值,中心像素差值卷积可以突出特征图中的边缘和纹理信息,从而提高目标区域的特征表示。

本文在 YOLOv8 网络的颈部添加 TFE 模块,并采用残差连接,将原始特征与抑制噪声后的特征相结合,突出目标特征的显示,减少背景干扰,达到提高检测精度的目的。

3 实验与分析

3.1 实验环境与数据集

本文实验基于 Windows10 专业版操作系统,运行内存 32 GB, CPU 为 Intel(R) Core(TM) i5-12490F 3.00 GH, GPU 为 NVIDIA GeForce RTX 4060Ti 16 GB。CUDA 版本为 12.2, Pytorch 版本为 2.1.0, 实验使用 Python 语言实现。实验训练参数设置如下:输入图像宽高为 608, Batch-Size 设为 8, 迭代次数为 200, 初始学习率为 0.01, 权重衰减为 0.0005, 动量为 0.937。

本文在暗光环境目标检测数据集 ExDark 上进行实验,对提出的方法进行评估。ExDark 数据集共有 7 363 张图像,按照 8 : 2 的比例划分训练集与测试集,训练集和测试集的图像数量分别为 5 884 张和 1 479 张。数据集总共标注了 12 个类别的目标,其中包括自行车、船、瓶子、公交车、汽车、猫、椅子、杯子、狗、摩托车、人以及桌子。

3.2 评价指标

实验结果以平均检测精度(mean average precision, mAP)、参数量(params)、计算量(GFLOPS)等指标衡量模型性能。mAP 的计算公式如下:

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

$$AP = \int_0^R P(R) dR \quad (8)$$

$$mAP = \frac{\sum_{i=1}^c AP_i}{C}$$

式中: TP (true positives) 表示被正确识别的正样本数量。 FP (false positives) 表示被错误识别为正样本的负样本数量。 FN (false negatives) 表示被错误识别为负样本的正样本数量。 P 表示准确率, R 表示召回率。 $P(R)$ 表示不

同召回率下的准确率,即召回率-精确率曲线。 AP_i 表示某一类别 i 的平均精度。 C 表示类别总数,对 C 个类别的 AP 取平均值,即得到平均精度 mAP 指标。本文采用 IoU 阈值为 0.5 时的 mAP 作为检测指标。

3.3 对比实验

本文使用目前主流的通用目标检测算法与针对低照度环境下的目标检测方法 MAET、Zero-DCE^[21] + YOLOv8s、Dark-YOLO^[27] 以及基于 YOLOv8n 改进的 DarkYOLOv8^[15] 进行对比实验。并在 ExDark 数据集上进行对比。

MAET 算法通过自监督学习方法对图像的光照变换以及噪声模型进行预测,将正常光照图像合成为低照度图像。然后将正常光照图像与合成低照度图像输入检测网络进行训练,提高训练效果。这种方法取得了精度的提升,但是图像变换模块设计复杂,导致模型计算量大大增加。Zero-DCE+YOLOv8s 算法通过 Zero-DEC 网络对低照度图像的亮度曲线进行预测,增强图像亮度。这种方法可以获得视觉效果上的正常光照图像,但与目标检测器所需要的特征数据存在偏差。Dark-YOLO 模型在检测网络的颈部使用空洞卷积增强感受野,实现不同尺度特征的加强融合,达到增强特征表达的目的。DarkYOLOv8 通过替换骨干网络增强算法的特征提取能力,以及在颈部添加动态特征融合注意力块增强多尺度特征的融合,取得了一定的效果,但是没有考虑背景噪声对特征信息的干扰。相比于上述方法,本文通过增强目标检测器所需要的特征信息并减少噪声对目标特征区域的干扰,获得了更好的检测性能。本文算法与上述算法的对比结果如表 1 所示。

表 1 对比实验

Table 1 Comparison experiment

方法	参数量/M	计算量	mAP/%
YOLOv8s	11.14	25.87	72.60
Faster-RCNN ^[4]	136.91	369.98	67.13
FCOS ^[28]	32.14	176.00	68.40
YOLOv3 ^[1]	61.97	67.84	73.19
YOLOv5s ^[29]	7.02	16.00	67.10
YOLOv10s ^[30]	8.07	24.80	71.50
MAET ^[19]	65.32	215.91	74.00
Zero-DCE ^[21] + YOLOv8s	11.15	33.32	71.91
Dark-YOLO ^[27]	—	—	74.76
DarkYOLOv8 ^[15]	—	8.53	70.10
本文改进+YOLOv8n	3.05	11.12	71.88
本文	11.18	29.59	75.63

由表中数据可知,本文所提方法在保持较低的参数量与计算量前提下,检测精度保持领先水平。与针对低照度

环境提出的目标检测方法 MAET、Zero-DCE+YOLOv8s 与 Dark-YOLO 相比,本文方法在较小计算量前提下, mAP 分别领先了 1.63%、3.72%、0.87%。由于 DarkYOLOv8 算法是在 YOLOv8n 的基础上改进而来,为保证对比结果的有效性,将本文所提出的改进方法应用在 YOLOv8n 网络中。检测精度相比于 DarkYOLOv8 提高了 1.78%。与主流通用目标检测方法 Faster-RCNN、FCOS^[28]、YOLOv3、YOLOv5s^[29]、YOLOv10s^[30] 以及本文的基准模型 YOLOv8s 相比,本文算法的检测精度具有领先优势。这些数据表明,本文提出的方法在低照度环境下的目标检测任务中,具有更好的性能,验证了本文算法的有效性。

3.4 消融实验

为了进一步验证本文所提出的各项改进方法的有效性,对各个改进方法进行了消融实验,实验结果如表 2 所示。

表 2 消融实验

Table 2 Ablation experiment

YOLOv8s	M_0	M_1	M_2	参数量/ M	计算量	mAP/ %
✓	—	—	—	11.14	25.87	72.60
✓	✓	—	—	11.17	29.57	73.02
✓	✓	✓	—	11.18	29.59	74.82
✓	—	—	✓	11.14	25.87	73.54
✓	✓	✓	✓	11.18	29.59	75.63

其中 M_0 表示使用 FEM 模块,但仅使用简单聚合方法处理不同尺度下的特征信息,而不使用 SFA 模块进行高效融合。 M_1 表示在 M_0 方法的基础上使用 SFA 模块对不同尺度特征进行高效融合。 M_2 表示在颈部添加 TFE 模块抑制噪声干扰。

由表 2 可知,原始模型 YOLOv8s 在 ExDark 数据集上获得了 72.60% 的检测精度,当引入简单特征聚合方法的图像特征增强模块后,检测精度提高了 0.42%,在此基础上加入高效的 SFA 模块融合特征,获得了 2.22% 的精度提升。实验结果表明,仅对不同尺度的特征信息进行简单聚合,不能有效地利用这些不同层次的特征信息,无法获得明显的精度提升。当加入 SFA 模块之后,获得了较大的精度提升,这说明本文提出的 FEM 模块能够有效地增强目标检测精度。同时,在颈部添加 TFE 模块后,精度提高了 0.94%。本文对添加 TFE 模块前后的颈部特征图进行了可视化操作,如图 5 所示,添加 TFE 模块后,特征图中的有效目标边缘更加清晰,并且,抑制了背景噪声干扰使得目标区域更加明显。最终改进方法与原始 YOLOv8s 模型相比,检测精度提高了 3.03%。

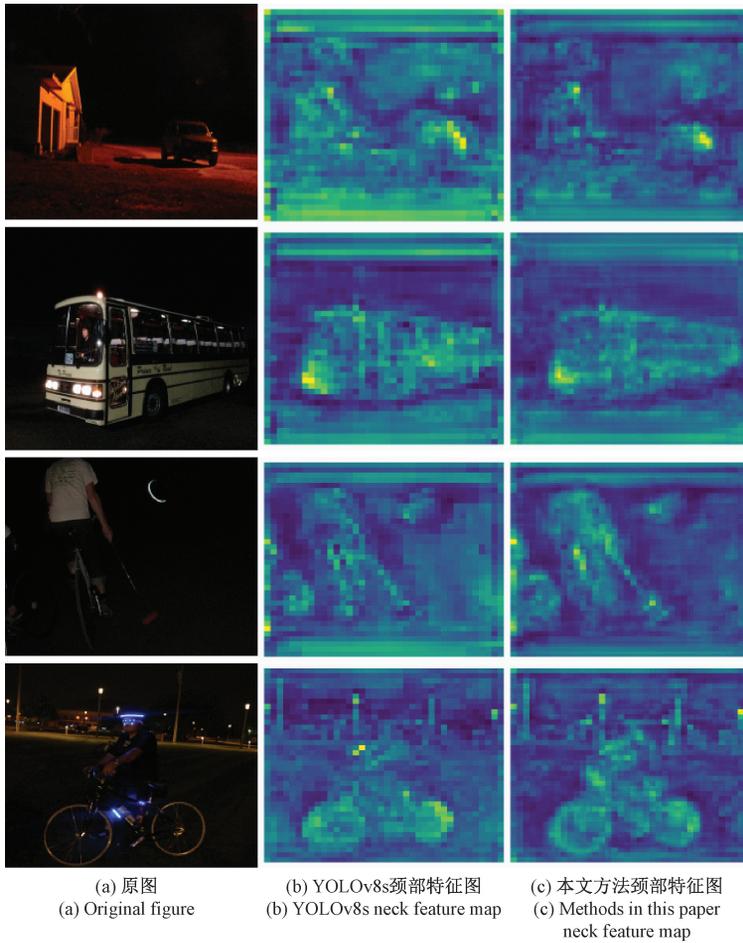


图5 特征图可视化

Fig. 5 Feature map visualization

4 结 论

针对低照度环境下的目标特征不明显导致检测任务性能下降的问题,本文对YOLOv8s模型进行改进,提出了一种基于特征增强的FE-YOLO算法。本文设计了图像特征增强模块,通过高效融合不同尺度下的图像特征信息,共同提高输入网络的图像质量。并将图像特征增强模块与目标检测网络级联,仅使用目标检测的损失值对整体网络进行优化,实现端到端的联合训练。此外,本文使用中心像素差值卷积设计了一个目标特征增强模块,通过背景信息的自相似性,抑制背景噪声的干扰,增强目标区域的特征表达,提高目标的检测精度。实验结果表明,与目前主流目标检测模型相比,FE-YOLOv8在低光照数据集ExDark上实现了更高的检测精度。未来将进一步探索本文算法在雾天、雨雪等不同退化场景中的检测效果。

参考文献

[1] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv preprint arXiv: 1804.02767, 2018.

[2] DIWAN T, ANIRUDH G, TEMBHURNE J V. Object detection using YOLO: Challenges, architectural successors, datasets and applications[J]. Multimedia Tools and Applications, 2023, 82(6): 9243-9275.

[3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.

[4] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. ArXiv preprint arXiv: 1506.01497, 2015.

[5] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [C]. IEEE International Conference on Computer Vision, 2017: 2961-2969.

[6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. Computer Vision-

- ECCV 2016, 2016: 21-37.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [8] SASAGAWA Y, NAGAHARA H. YOLO in the dark-domain adaptation method for merging multiple models [C]. Computer Vision-ECCV 2020, 2020: 345-359.
- [9] LI C C, DU D W, ZHANG L B, et al. Spatial attention pyramid network for unsupervised domain adaptation[C]. Computer Vision-ECCV 2020, 2020: 481-497.
- [10] CUI Z T, LI K CH, GU L, et al. You only need 90k parameters to adapt light: A light weight transformer for image enhancement and exposure correction[J]. ArXiv preprint arXiv:2205.14871, 2022.
- [11] LIU W Y, REN G F, YU R SH, et al. Image-adaptive YOLO for object detection in adverse weather conditions [C]. AAAI Conference on Artificial Intelligence, 2022, 36(2): 1792-1800.
- [12] KALWAR S, PATEL D, AANEGOLA A, et al. Gdip: Gated differentiable image processing for object detection in adverse conditions [C]. 2023 IEEE International Conference on Robotics and Automation (ICRA), 2023: 7083-7089.
- [13] 王非凡, 陈希爱, 任卫红, 等. 基于图像自适应增强的低照度目标检测算法[J]. 计算机工程. DOI: 10.19678/j.issn.1000-3428.0068407.
WANG F F, CHEN X AI, REN W H, et al. A Low Illumination Object Detection Algorithm Based on Image Adaptive Enhancement [J]. Computer Engineering. DOI: 10.19678/j.issn.1000-3428.0068407.
- [14] 杜运亮, 王明甲. 基于半监督域适应的微弱光环境下行人检测研究[J]. 电子测量与仪器学报, 2024, 38(1): 106-113.
DU Y L, WANG M J. Research on pedestrian detection in low-light conditions based on semi-supervised domain adaptation [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(1):106-113.
- [15] 蔡腾, 陈慈发, 董方敏. 结合 Transformer 和动态特征融合的低照度目标检测[J]. 计算机工程与应用, 2024, 60(9): 135-141.
CAI T, CHEN C F, DONG F M. Low-light object detection combining transformer and dynamic feature fusion[J]. Computer Engineering and Applications, 2024, 60(9): 135-141.
- [16] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [17] 李昶昱, 葛磊. 基于 YOLOv7 的轻量级低照度目标检测算法[J]. 激光与光电子学进展, 2024, 61(14):1437004.
LI CH Y, GE L. Lightweight low-light target detection algorithm based on YOLOv7[J]. Laser & Optoelectronics Progress, 2024, 61(14):1437004.
- [18] YIN X CH, YU ZH D, FEI Z T, et al. PE-YOLO: Pyramid enhancement network for dark object detection[C]. International Conference on Artificial Neural Networks. Cham: Springer Nature Switzerland, 2023: 163-174.
- [19] CUI Z T, QI G J, GU L, et al. Multitask aet with orthogonal tangent regularity for dark object detection[C]. IEEE/CVF International Conference on Computer Vision, 2021: 2553-2562.
- [20] 冯欣, 王思平, 张智先, 等. 一种面向机器视觉感知的暗光图像增强网络[J]. 计算机应用研究, 2024, 41(6): 1910-1915.
FENG X, WANG S P, ZHANG ZH X, et al. Dark light image enhancement network for machine vision perception[J]. Application Research of Computers, 2024, 41(6): 1910-1915.
- [21] LI CH Y, GUO CH L, LOY CH CH. Learning to enhance low-light image via zero-reference deep curve estimation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(8): 4225-4238.
- [22] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017, 30. DOI: 10.48550.
- [23] 张蕊, 高诗博, 赵霞, 等. 基于改进 YOLOv5s 的无人驾驶夜间车辆目标检测算法[J]. 电子测量技术, 2023, 46(17):87-93.
ZHANG R, GAO SH B, ZHAO X, et al. Algorithm on nighttime target detection for unmanned vehicles based on an improved YOLOv5s [J]. Electronic Measurement Technology, 2023,46(17):87-93.
- [24] 邹红波, 陈俊廷, 柴延辉, 等. 基于改进的 YOLOv7 的雾天场景下绝缘子及其自爆缺陷检测方法[J]. 国外电子测量技术, 2023(10):1-11.
ZOU H B, CHEN J T, CHAI Y H, et al. Detection method of insulator and self-explosion defect in foggy scene based on improved YOLOv7 [J]. Foreign Electronic Measurement Technology, 2023(10):1-11.
- [25] WANG W J, YANG W H, LIU J Y. H1a-face: Joint

- high-low adaptation for low light face detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 16195-16204.
- [26] SU ZH, LIU W ZH, YU Z T, et al. Pixel difference networks for efficient edge detection[C]. IEEE/CVF International Conference on Computer Vision, 2021: 5117-5127.
- [27] 江泽涛, 肖芸, 张少钦, 等. 基于 Dark-YOLO 的低照度目标检测方法[J]. 计算机辅助设计与图形学学报, 2023, 35(3): 441-451.
- JIANG Z T, XIAO Y, ZHANG SH Q, et al. Low-illumination object detection method based on Dark-YOLO[J]. Journal of Computer-Aided Design & Computer Graphics, 2023, 35(3): 441-451.
- [28] TIAN ZH, SHEN CH H, CHEN H, et al. FCOS: Fully convolutional one-stage object detection[J]. ArXiv preprint arXiv:1904.01355, 1904.
- [29] WANG J, YANG P, LIU Y SH, et al. Research on improved YOLOv5 for low-light environment object detection[J]. Electronics, 2023, 12(14): 3089.
- [30] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458, 2024.

作者简介

黄玉龙, 硕士研究生, 主要研究方向为图像增强、目标检测等。

E-mail: hyl3416@163.com

张晓玲(通信作者), 博士, 副教授, 主要研究方向为图像处理、视觉检测、机器人视觉导航等。

E-mail: 23175809@qq.com